

A probabilistic construction of small complete caps in projective spaces

Daniele Bartoli¹, Stefano Marcugini, Fernanda Pambianco

Department of Mathematics and Computer Sciences of Perugia University, Via Vanvitelli 1, Perugia, 06123

Abstract

In this work complete caps in $PG(N, q)$ of size $O(q^{\frac{N-1}{2}} \log^{300} q)$ are obtained by probabilistic methods. This gives an upper bound asymptotically very close to the trivial lower bound $\sqrt{2}q^{\frac{N-1}{2}}$ and it improves the best known bound in the literature for small complete caps in projective spaces of any dimension. The result obtained in the paper also gives a new upper bound for $l(m, 2, q)_4$, that is the minimal length n for which there exists an $[n, n - m, 4]_q$ covering code with given m and q .

Keywords: Complete caps, Projective spaces, Quasi-perfect codes, Covering codes

2010 MSC: 51E22, 51E20

1. Introduction

A n -cap in an (affine or projective) Galois space over the finite field with q elements \mathbb{F}_q is a set of n points no three of which are collinear. A n -cap is said to be complete if it is not contained in a $(n + 1)$ -cap. A plane n -cap is also called a n -arc.

A central problem on caps is determining the spectrum of the possible sizes of complete caps in a given space, see the survey paper [1] and the references therein. As mentioned above, of particular interest for applications to Coding Theory is the lower part of the spectrum.

For the size of the smallest complete cap in the projective space $PG(N, q)$ of dimension N over \mathbb{F}_q , the trivial lower bound is

$$\sqrt{2}q^{\frac{N-1}{2}}. \quad (1)$$

General constructions of complete caps whose size is close to this lower bound are only known for q even and N odd; see [2, 3, 4, 5]. When N is even, complete caps

Email addresses: daniele.bartoli@dmf.unipg.it (Daniele Bartoli), gino@dmf.unipg.it (Stefano Marcugini), fernanda@dmf.unipg.it (Fernanda Pambianco)

¹Corresponding author

of size of the same order of magnitude as $cq^{N/2}$, with c a constant independent of q , are known to exist for both the odd and the even order case, see [3, 6, 4, 7, 8].

If q is odd and $N \equiv 0 \pmod{4}$, small complete caps can be obtained via the product method for caps from bicovering plane arcs. It has been shown in [7] that the cartesian product of a bicovering k -arc A in the affine plane $AG(2, q)$ and the cap of size $q^{\frac{N-2}{2}}$ in the affine space $AG(N-2, q)$ arising from the blow-up of a parabola of $AG(2, q^{(N-2)/2})$ is a complete cap in $AG(N, q)$. Via the natural embedding of $AG(N, q)$ in $PG(N, q)$ it is possible to obtain from a complete cap in $AG(N, q)$ a complete cap in $PG(N, q)$ of the same magnitude.

In [9] the authors obtain caps of size $(k+1)q^{\frac{N-2}{2}}$ in $AG(N, q)$ when the k -arc A is almost bicovering, that is, a complete arc which bicovers all points in $AG(2, q) \setminus A$ but one.

By similar methods, in [10] the authors provide new complete caps in $AG(N, q)$ with roughly $q^{(4N-1)/8}$ points, studying both plane cubics with a node and plane cubics with an isolated double point.

In [9] the existence of complete caps in $AG(N, q)$, $N \equiv 0 \pmod{4}$, of size of the same order of magnitude as $2pq^{\frac{1}{2}(N-\frac{1}{4})}$, provided that the characteristic p of \mathbb{F}_q is large enough and $\log_p q > 8$, is established.

The exact value $t_2(N, q)$ of the minimum size of a complete cap in $PG(N, q)$ is known only for few pairs (N, q) : for instance in the case $N = 3$, $t_2(3, q)$ is known only for $q \leq 7$; see [11, Table 3].

In the case $N = 3$ according to the survey paper [1], the smallest known complete caps in $PG(3, q)$, with q arbitrary large, have size approximately $q^{3/2}/2$ and were presented by Pellegrino in 1998 [12]. However, Pellegrino's completeness proof appears to present a major gap, and counterexamples can be found; see [13, Section 2].

In this work, the existence of complete caps of size

$$O(q^{\frac{N-1}{2}} \log^{300} q) \quad (2)$$

in projective spaces $PG(N, q)$, with $N \geq 3$, is established by probabilistic methods. This gives an upper bound asymptotically very close to the trivial lower bound (1) and it improves the best known bound in the literature for small complete caps in projective spaces of any dimension.

In this paper we use techniques similar to those of [14] to extend the results from complete arcs in projective plane to complete caps in any projective space of dimension greater than two. In [14] the authors proved the following theorem; see [14, Theorem 1.1].

Theorem 1.1. *There are positive constants c and M such that the following holds. In every projective plane of order $q \geq M$, there is a complete arc of size at most*

$$q^{1/2} \log^c q.$$

They used new and powerful concentration results (see [15]) and in particular they proved this stronger result; see [14, Theorem 1.2].

Theorem 1.2. *There are absolute constants c and M such that in any projective plane of order $q \geq M$, one can find an arc with $\Theta(q^{1/2} \log^{1/2} q)$ points whose secants cover all but $q^{1/2} \log^c q$ points of the plane.*

Also, they provided an efficient randomized algorithm which produces the desired arc with probability close to 1 and runs in $\Theta(\log^{5/2} q)$ steps, where each step consist of $O(q^4)$ basic operations; see [14, Section 2.2].

The main result of this work, in term of complete caps in any projective space (N, q) is the following.

Theorem 1.3. *There are positive constants c and M such that in every projective space of order $q \geq M$ and dimension N , there is a complete cap of size at most*

$$q^{\frac{N-1}{2}} \log^c q.$$

Also, as in [14], a randomized algorithm to construct the desired complete caps can be easily deduced.

Let \mathbb{F}_q be the finite field with q elements. A q -ary linear code \mathcal{C} of length n and dimension k is a k -dimensional linear subspace of \mathbb{F}_q^n . The Hamming weight $w(\mathbf{x})$ of \mathbf{x} is the number of nonzero positions in a vector $\mathbf{x} \in \mathcal{C}$. The minimum distance of \mathcal{C} is defined as

$$d(\mathcal{C}) := \min\{w(\mathbf{x}) | \mathbf{x} \in \mathcal{C}, \mathbf{x} \neq 0\}$$

and a q -ary linear code of length n , dimension k , and minimum distance d is denoted as an $[n, k, d]_q$ -code. An $[n, k, d]_q$ -code can correct at most t -errors, where $t = \lfloor \frac{d-1}{2} \rfloor$. The covering radius of \mathcal{C} is the minimum integer $R(\mathcal{C})$ such that for any vector $v \in \mathbb{F}_q^n$ there exists $x \in \mathcal{C}$ with $w(v - x) \leq R$. An $[n, k, d]_q$ -code with covering radius R is denoted by $[n, k, d]_q R$. If $R = t$ then \mathcal{C} is said to be perfect. As there are only finitely many classes of linear perfect codes, of particular interest are those codes \mathcal{C} with $R = t + 1$, called quasi-perfect codes; see [16, 17, 18]. The covering density $\mu(\mathcal{C})$, introduced in [19], is one of the parameters characterizing the covering quality of an $[n, k, d]_q R$ -code \mathcal{C} and it is defined by

$$\mu(\mathcal{C}) = \frac{1}{q^{n-k}} \sum_{i=0}^R (q-1)^i \binom{n}{i}.$$

Note that $\mu(\mathcal{C}) \geq 1$, and that equality holds when \mathcal{C} is perfect. Clearly, among codes with the same codimension and covering radius the shortest ones have the best covering density. Therefore the problem of determining the minimal length n for which there exists an $[n, n-m, d]_q R$ -code with given m, q, d , and R , has been broadly investigated, see [20]. Throughout, such minimal length will be denoted as $l(m, R, q)_d$.

In the case $R = 2$ and $d = 4$, that is, for quasi-perfect linear codes that are both 1-error correcting and 2-error detecting, the columns of a parity check matrix of an $[n, n-m, 4]_q$ -code can be considered as points of a complete n -cap in the finite projective space $PG(m-1, q)$. For this reason these codes have

been investigated also through their connection with Projective Geometry; see e.g. [8].

The construction of complete caps presented in this paper gives, via the connection between Coding Theory and Projective Geometry, the following upper bound on $l(N + 1, 2, q)_4$.

Theorem 1.4. *There exists a positive constants M such that for every $q \geq M$*

$$l(N + 1, 2, q)_4 = O(q^{\frac{N-1}{2}} \log^{300} q).$$

The paper is organized as follows. In Section 2 the algorithm based on the nibble method for the construction of the complete caps is described. In Section 3 the notations used in this work are presented. In Section 4.1 the main features of the two phases of the algorithm are listed and the Main Theorem is proven via these properties. In Section 6 all the principal probabilistic instruments are collected. In Sections 7 and 8 all the properties used in the proof of the main Theorem are proven.

2. Algorithm

2.1. The nibble method

The algorithm used to construct complete caps is the same as in [14]. This algorithm is based on the nibble method and in the following we recall the main features of this method. For a more detailed introduction see [14, Section 2.1].

Usually to construct a cap one can proceed selecting the points one by one (for instance see [21, 22]) adding at each step a point not lying on bisecants of the cap at the present step. The choice can be done in a complete random way, taking the first point in a certain order (see [22]), or selecting the point which maximizes an objective function (see [21]). Usually the results obtained in this way are good if we are interested in small complete caps; see [21, 22, 23, 24, 25] and the reference therein.

The nibble method, or dynamic random construction using nibble (DRC), is an approximated version of the previous algorithm. Instead of selecting one element at each step, DRC randomly and independently chooses points, not selected yet, with a certain probability so that a bunch of elements are chosen together. This is called a “nibble” and its size is the number of chosen elements or its expectation. The first problem, when selecting a set of points instead of a single point at each step, is that the new set obtained by the union of the nibble and the cap at the previous step could be not a cap. In order to avoid this, only a subset of the nibble is added to the cap. Usually, each point in the nibble (we will refer to the points in the nibble as *chosen* points) which does not lie on bisecants of the cap or of the nibble itself is *selected*, i.e. is chosen to be added to the constructing cap. As not all the chosen points are selected, some points could unnecessarily be eliminated, but in this way the set of remaining elements available for the subsequent steps can be well-understood. The size of a nibble cannot be too big, since in this case many of the chosen points contribute to

at least one conflict (that is lies on some bisecants), and it would be hard to predict the structure of the cap at each step and/or too many elements would unnecessarily be discarded. Therefore the size of the nibble should be small enough to have that most of the chosen elements do not cause any conflict and only few elements would be unnecessarily discarded. For instance, if we choose $\theta q^{\frac{N-1}{2}}$ random points from a projective space of order q , each chosen point causes a conflict with probability at most

$$(q^{N-1} + q^{N-2} + \dots + q + 1) \binom{q}{2} \left(\theta q^{-\frac{N+1}{2}} \right)^2 \leq \theta^2.$$

If θ is small enough, then most of the chosen points do not cause any conflict.

2.2. The algorithm

The cap is constructed in the following way. At the beginning, the starting cap A_0 is empty. Let $\Omega_0 = S_0$ be the set of all the points of the projective space $PG(N, q)$. Roughly speaking, at each step Ω_i is essentially the set of points which are not covered by the cap A_i , while S_i is a subset of Ω_i . At the each step i , a random subset $B_i \subset S_i$ is selected, choosing each point from S_i independently with the same probability p_i . The set B_i is the *nibble* and only a subset of B_i is added to A_i to obtain the new cap A_{i+1} . This subset $M_i \subset B_i$ is the set of the points not causing any conflict. At the subsequent step, Ω_{i+1} is obtained from Ω_i by deleting all the points covered by the secants of A_{i+1} or in B_i , while S_{i+1} is obtained by deleting from S_i the points covered by the secants of A_{i+1} or in B_i plus a few more points, chosen randomly: in this way certain structural properties of the S_i 's are preserved. The process is repeated until all but $q^{\frac{N-1}{2}} \log^c q$ points are covered by the secants of the current arc. In the following we set $\theta = \log^{-2} q$.

The algorithm acts as follows. At each step we use three different “sub-phases”: **choose**, **delete**, and **compensate**.

Start : Ω_0, S_0 are both $PG(N, q)$, $A_0 = \emptyset$. We also consider the quantities a_i and b_i . At the beginning $a_0 = 0$ and $b_0 = 1$, while at the i -th step

$$a_i = \frac{|A_i|}{q^{\frac{N-1}{2}}} \quad \text{and} \quad b_i = \frac{|S_i|}{q^N + q^{N-1} + \dots + q + 1}.$$

Choose : At each step a point v in S_i is chosen with probability $p_i = \theta(b_i q^{\frac{N+1}{2}})^{-1}$. The set of all the chosen points is B_i . A point x in B_i is *good* in $A_i \cup B_i$ if there are no two points in $A_i \cup B_i$ collinear with x . The set M_i is the set of all the good points. So $A_{i+1} = A_i \cup M_i$.

Delete : Delete from Ω_i all the points in bisecants of A_{i+1} or in B_i . Let D_i be the set of deleted points and, if $v \in \Omega_i$, let $P_i(v) = Pr(v \in D_i)$. Let P_i^u and P_i^l be the upper and the lower bounds for these probabilities.

Compensate : S_{i+1} is obtained from S_i deleting the points of D_i and independently the points of S_i with probability

$$P_i^{com}(v) = \frac{P_i^u - P_i(v)}{1 - P_i(v)}.$$

Let R_i be the set of the removed points, then

$$\Omega_{i+1} = \Omega_i \setminus D_i, \quad S_{i+1} = S_i \setminus (D_i \cup R_i), \quad A_{i+1} = A_i \cup M_i$$

and

$$b_{i+1} = b_i(1 - P_i^u) = \prod_{j=1}^i (1 - P_j^u), \quad b'_{i+1} = b'_i(1 - P_i^\ell) = \prod_{j=1}^i (1 - P_j^\ell).$$

Stop : The algorithm stops after K steps if K is the first integer such that

$$b_K \leq q^{-\frac{N+1}{2}} \log^c q,$$

for some constant c (we will set $c = 300$, as in [14]).

Remark 2.1. *All the properties shown in the following are proven for all $i \leq o(\log^3 q)$.*

The importance of the sub-phase of Compensation is explained in the following remark.

Remark 2.2. *The operation of “compensation” is made in order to give the same probability to the points in S_i to be in S_{i+1} . In fact, if $p = P_i(v)$, then*

$$P(v \notin S_{i+1} | v \in S_i) = p + (1 - p) \frac{P_i^u - p}{1 - p} = P_i^u.$$

So,

$$\mathbb{E}(|S_{i+1}|) = |S_i|(1 - P_i^u).$$

In order to prove the properties on the nibble and on the cap at every step, we divide the search into two phases. The first phase continues when

$$b_i \geq \frac{\log^{c_1} q}{q},$$

while the second phase continues when

$$\frac{\log^c q}{q^{\frac{N+1}{2}}} \leq b_i \leq \frac{\log^{c_1} q}{q}.$$

3. Notations

Let ℓ be a line and u, v points of $PG(N, q)$. Also, if a, b, c are points of the space, then $[abc]$ indicates that the three points are collinear, while (ab) is the line through a and b . In the following we will use the following sets.

- $A_i(v) = \{x \in S_i \setminus \{v\} | \exists u \in A_i : [xuv]\}$ is the set of all the points of S_i belonging to the cone of lines through $v \in \Omega_i$ and the points of A_i ;
- $A'_i(v) = \{x \in \Omega_i \setminus \{v\} | \exists u \in A_i : [xuv]\}$ is the set of all the points of Ω_i belonging to the cone of lines through $v \in \Omega_i$ and the points of A_i ;
- $A_i(u_1, \dots, u_m) = \bigcap_{j=1}^m A_i(u_j)$;

- $B_i(v) = \{x \in S_i \setminus \{v\} \mid \exists u \in B_i : [xuv]\}$ is the set of all the points of S_i belonging to the cone of lines through $v \in \Omega_i$ and the points of B_i ; $B_i(u_1, \dots, u_m) = \bigcap_{j=1}^m B_i(u_j)$;
- $S_i(\ell) = S_i \cap \ell$: note that if ℓ is a bisecant of the cap then $S_i(\ell)$ is empty; $S_i(\ell, u) = S_i(\ell) \cap A_i(u)$; $S_i(\ell, u, v) = S_i(\ell) \cap A_i(u) \cap A_i(v)$;
- $\Omega_i(\ell) = \Omega_i \cap \ell$; $\Omega_i(\ell, u) = \Omega_i(\ell) \cap A'_i(u)$; $\Omega_i(\ell, u, v) = \Omega_i(\ell) \cap A'_i(u) \cap A'_i(v)$;
- $T_i(v) = \{\{x, y\} \mid x, y \in S_i \setminus \{v\}, [vxy]\}$, the set of the unordered pairs of points of S_i collinear with v ; $T'_i(v) = \{(x, y) \mid x, y \in S_i \setminus \{v\}, [vxy]\}$, the set of the ordered pairs of points of S_i collinear with v .

From the above definitions the following hold:

$$|T_i(v)| = \sum_{\ell \ni v} \binom{|S_i(\ell)| - 1}{2}; \quad |S_i| = \sum_{\ell} \frac{|S_i(\ell)|}{q^{N-1} + \dots + q + 1};$$

$$|\Omega_i| = \sum_{\ell} \frac{|\Omega_i(\ell)|}{q^{N-1} + \dots + q + 1}; \quad |A_i(v)| = \sum_{\substack{\ell \ni a, v \\ a \in A_i}} (|S_i(\ell)| - 1).$$

Also, given two functions $f(q), g(q)$, as usual, we say that $g(q) = o(f(q))$, $g(q) = O(f(q))$, $g(q) = \Theta(f(q))$, if

$$\lim_{q \rightarrow +\infty} \frac{g(q)}{f(q)} = 0, \quad \lim_{q \rightarrow +\infty} \frac{g(q)}{f(q)} \leq K \in \mathbb{N}, \quad \lim_{q \rightarrow +\infty} \frac{g(q)}{f(q)} = K \in \mathbb{N},$$

respectively. In the following we say that an event A “occurs with very high probability” if $Pr(\overline{A}) \leq e^{-w \log q}$ for some constant $w \in \mathbb{N}$. We consider in all the proofs that the dimension N of the space is *greater than 2*.

4. Main Lemma

In the following lemma all the properties which hold at each step are summarized.

Lemma 4.1. *The following holds for any $i \leq o(\log^3 q)$.*

- **Phase 1**

- **Primary Properties**

$$\theta q^{\frac{N-1}{2}}(1 - o(1)) \leq |M_i| \leq \theta q^{\frac{N-1}{2}}(1 + o(1)) \quad |B_i| \leq 2\theta q^{\frac{N-1}{2}} \quad (3)$$

$$b_{i+1}q(1 - (i+1)\log^{-13}q) \leq |S_{i+1}(\ell)| \leq b_{i+1}q(1 + (i+1)\log^{-13}q) \quad (4)$$

$$|\Omega_{i+1}(\ell)| \leq b'_i q(1 + (i+1)\log^{-13}q) \quad (5)$$

– **Secondary Properties**

$$|S_{i+1}(\ell, v)| \leq 8(i+1)a_{i+1}b_{i+1}\sqrt{q} + (i+1)\log^{40} q \quad (6)$$

$$|S_{i+1}(\ell, u, v)| \leq (i+1)\log^4 q \quad (7)$$

$$|A_{i+1}(u, v)| \leq (i+1)b_{i+1}q + (i+1)\log^{40} q \quad (8)$$

$$|\Omega_{i+1}(\ell, v)| \leq 8(i+1)a_{i+1}b'_{i+1}\sqrt{q} + (i+1)\log^{40} q \quad (9)$$

$$|\Omega_{i+1}(\ell, u, v)| \leq (i+1)\log^4 q \quad (10)$$

• **Phase 2**

$$\theta q^{\frac{N-1}{2}}(1 - o(1)) \leq |M_i| \leq \theta q^{\frac{N-1}{2}}(1 + o(1)) \quad |B_i| \leq 2\theta q^{\frac{N-1}{2}} \quad (11)$$

$$\frac{1}{2}b_{i+1}^2 q^{N+1}(1 - 3(i+1)\log^{-13} q) \leq |T_{i+1}(v)| \leq \frac{1}{2}b_{i+1}^2 q^{N+1}(1 + 3(i+1)\log^{-13} q) \quad (12)$$

$$a_{i+1}b_{i+1}q^{\frac{N+1}{2}}(1 - O(i\theta^2)) \leq |A_{i+1}(v)| \leq a_{i+1}b_{i+1}q^{\frac{N+1}{2}}(1 + O(i\theta^2)) \quad (13)$$

$$(1 - \log^{-10} q)^{i+1}b_{i+1}q^N \leq |S_{i+1}| \leq (1 + \log^{-10} q)^{i+1}b_{i+1}q^N \quad (14)$$

$$|\Omega_{i+1}| \leq (1 + \log^{-10} q)^{i+1}q^N b'_{i+1} \quad (15)$$

Remark 4.2. From the properties of the first phase, recalling that $p_i = \frac{\theta}{b_i q^{\frac{N+1}{2}}}$, we have

$$|A_i(v)| = |A_i||S_i(\ell)|, \quad |T_i(v)| \simeq \frac{1}{2}(q^{N-1} + q^{N-2} + \dots + q + 1)|S_i(\ell)|^2$$

and therefore

$$\frac{1}{2}b_{i+1}^2 q^{N+1}(1 - 3(i+1)\log^{-13} q) \leq |T_{i+1}(v)| \leq \frac{1}{2}b_{i+1}^2 q^{N+1}(1 + 3(i+1)\log^{-13} q) \quad (16)$$

$$a_{i+1}b_{i+1}q^{\frac{N+1}{2}}(1 - 2(i+1)\log^{-13} q) \leq |A_{i+1}(v)| \leq a_{i+1}b_{i+1}q^{\frac{N+1}{2}}(1 + 2(i+1)\log^{-13} q) \quad (17)$$

$$(1 - \log^{-10} q)^{i+1}b_{i+1}q^N \leq |S_{i+1}| \leq (1 + \log^{-10} q)^{i+1}b_{i+1}q^N \quad (18)$$

$$|\Omega_{i+1}| \leq (1 + \log^{-10} q)^{i+1}q^N b'_{i+1} \quad (19)$$

5. Proof of the Main Theorem via the Main Lemma

In this section we prove the Main Theorem using the properties listed in the Main Lemma. The proofs of Lemmas 5.1 and 5.2 are the same as [14, Lemma 3.2] and [14, Lemma 3.3] and therefore omitted.

Lemma 5.1. *We have*

$$P_i^u \leq p_i + p_i \max_v |A_i(v)| + p_i^2 \max_v |T_i(v)|,$$

where the maximum is taken over the set Ω_i .

The following estimates hold:

$$\begin{aligned} p_i &= \frac{\theta}{b_i q^{\frac{N+1}{2}}} \leq \frac{\theta}{\log^c q} = \log^{-302} q \leq \log^{-4} q. \\ p_i^2 \max_v |T_i(v)| &\simeq \frac{\theta^2}{b_i^2 q^{N+1}} b_i^2 q^{N+1} = \theta^2 = \log^{-4} q. \\ \log^{-4} q < p_i \max_v |A_i(v)| &\simeq \frac{\theta}{b_i q^{\frac{N+1}{2}}} i b_i \theta q^{\frac{N+1}{2}} = i \theta^2 \leq \log^{-1} q. \end{aligned}$$

Lemma 5.2. *We have*

$$P_i^l \geq p_i \min_v |A_i(v)| - 2p_i^2 \max_v |A_i(v)|^2 - p_i^3 \max_v |A_i(v)| \max_v |T_i(v)|,$$

where the maximum is taken over the set Ω_i .

The following estimates hold:

$$\begin{aligned} p_i \min_v |A_i(v)| &\simeq \frac{\theta}{b_i q^{\frac{N+1}{2}}} i b_i \theta q^{\frac{N+1}{2}} = i \theta^2. \\ 2p_i^2 \max_v |A_i(v)|^2 &\simeq 2i^2 \theta^4 = o(i \theta^2). \\ p_i^3 \max_v |A_i(v)| \max_v |T_i(v)| &\simeq i \theta^2 \frac{\theta^2}{b_i^2 q^{N+1}} b_i^2 q^{N+1} = i \theta^4 = o(i \theta^2). \end{aligned}$$

Let K be the first integer such that

$$b_K < \frac{\log^c q}{q^{\frac{N+1}{2}}},$$

that is K is the total number of steps of the algorithm.

Lemma 5.3. *Suppose that Lemma 4.1 holds. Then*

$$K = \Theta(\theta^{-1} \log^{1/2} q) = \Theta(\log^{5/2} q).$$

Proof. By the definition of K

$$\prod_{i=0}^{K-1} (1 - P_i^u) = b_K \leq q^{-\frac{N+1}{2}} \log^c q \leq b_{K-1} = \prod_{i=0}^{K-2} (1 - P_i^u).$$

Suppose $K \geq L = \lceil \sqrt{6N} \log^{5/2} q \rceil$, then

$$b_L \geq b_{K-1} \geq q^{-\frac{N+1}{2}} \log^c q \tag{20}$$

and

$$b_L = \prod_{i=1}^{L-1} (1 - P_i^u) \leq \prod_{i=1}^{L-1} (1 - P_i^\ell).$$

From (20) we have

$$\begin{aligned} -\frac{N+1}{2} \log q + c \log \log q &= \left(-\frac{N+1}{2} + \frac{\log \log q}{\log q} \right) \log q = \\ &= \left(-\frac{N+1}{2} + o(1) \right) \log q \leq \sum_{i=1}^{L-1} \log(1 - P_i^\ell). \end{aligned}$$

From Lemma 5.2 we have

$$P_i^\ell \geq \frac{1}{2} i \theta^2,$$

so

$$\log(1 - P_i^\ell) \leq \log \left(1 - \frac{1}{2} i \theta^2 \right) \leq -\frac{1}{3} i \theta^2.$$

Then

$$\sum_{i=1}^{L-1} \log(1 - P_i^\ell) \leq -\frac{1}{3} \sum_{i=1}^{L-1} i \theta^2 \leq -\frac{1}{7} L^2 \theta^2 \leq -\frac{6}{7} N \log q$$

which is impossible, then $K < L = \lfloor \sqrt{6N} \log^{5/2} q \rfloor$. Suppose that $K \leq L = \lfloor \sqrt{\frac{N}{6}} \log^{5/2} q \rfloor$, then

$$\prod_{i=1}^L (1 - P_i^u) = b_{L+1} \leq b_K \leq q^{-\frac{N+1}{2}} \log^c q \quad (21)$$

and then

$$\sum_{i=1}^L \log(1 - P_i^u) \leq \left(-\frac{K+1}{2} + o(1) \right) \log q$$

From Lemma 5.1 we have $P_i^u \leq 2i\theta^2$. Then

$$\log(1 - P_i^\ell) \geq \log(1 - 2i\theta^2) \geq -3i\theta^2.$$

$$-\frac{K}{2} \log q \leq \sum_{i=1}^L \log(1 - P_i^u) \leq \left(-\frac{K+1}{2} + o(1) \right) \log q,$$

which is impossible, therefore $K > \sqrt{\frac{N}{6}} \log^{5/2} q$ and $K = \Theta(\log^{5/2} q)$. \square

Corollary 5.4. *Suppose that Lemma 4.1 holds, then*

$$|A_K| = \Theta(q^{\frac{N-1}{2}} \log^{1/2} q).$$

Proof. By Lemma 5.3 the number of steps is $\Theta(\log^{5/2} q)$; since $|A_K| = \sum_{i=1}^K |M_i|$, then

$$|A_K| = \theta q^{\frac{N-1}{2}} \Theta(\log^{5/2} q) = \Theta(q^{\frac{N-1}{2}} \log^{1/2} q).$$

\square

Lemma 5.5. *Suppose that Lemma 4.1 holds, then for every i*

$$P_i^u - P_i^\ell = O(\log^{-3} q).$$

Proof. By Lemmas 5.1 and 5.2 we have

$$\begin{aligned} P_i^u - P_i^\ell &\leq p_i(1 + \max_v |A_i(v)| - \min_v |A_i(v)|) + p_i^2 \max_v |T_i(v)| + \\ &\quad + 2p_i^2 \max_v |A_i(v)|^2 + p_i^3 \max_v |A_i(v)| \max_v |T_i(v)|. \end{aligned}$$

The following estimations hold:

$$\begin{aligned} p_i^3 \max_v |A_i(v)| \max_v |T_i(v)| &= O(i\theta^4) = O(\log^{-11/2} q); \\ 2p_i^2 \max_v |A_i(v)|^2 &= O(2i^2\theta^4) = O(\log^{-3} q); \\ p_i^2 \max_v |T_i(v)| &= O(\theta^2) = O(\log^{-4} q). \end{aligned}$$

By Property (13)

$$\max_v |A_i(v)| - \min_v |A_i(v)| = O(i\theta^2 a_{i+1} b_{i+1} q^{\frac{N+1}{2}}),$$

then

$$p_i(\max_v |A_i(v)| - \min_v |A_i(v)|) = O(i^2\theta^4) = O(\theta^2 \log q) = O(\log^{-3} q),$$

and finally

$$p_i = O(\log^{-2-c} q).$$

□

Using the previous lemmas we are able to prove the main theorem.

Theorem 5.6 (Main Theorem). *There are positive constants c and M such that in every projective space of order $q \geq M$ and dimension n , there is a complete cap of size at most*

$$q^{\frac{N-1}{2}} \log^c q.$$

Proof. Suppose that Lemma 4.1 holds. By Corollary 5.4, $|A_N| = \Theta(q^{\frac{N-1}{2}} \log^{1/2} q)$. By (19)

$$|\Omega_K| \leq (1 + \log^{-10} q)^K q^N b'_K = (1 + o(1)) q^N \prod_{i=0}^{K-1} (1 - P_i^\ell)$$

and

$$|S_K| = \Theta(b_K q^N).$$

Then

$$\frac{|\Omega_K|}{|S_K|} = O\left(\prod_{i=0}^{K-1} \frac{1 - P_i^\ell}{1 - P_i^u}\right).$$

By Lemma 5.5 we have that

$$\log \left(\frac{1 - P_i^\ell}{1 - P_i^u} \right) = O(P_i^u - P_i^\ell) = O(\log^{-3} q).$$

Then

$$\log \frac{|\Omega_K|}{|S_K|} = O \left(\sum_{i=0}^{K-1} \log \left(\frac{1 - P_i^\ell}{1 - P_i^u} \right) \right) = O(K \log^{-3} q) = O(\log^{-1/2} q) = o(1),$$

and

$$|\Omega_K| = e^{o(1)} |S_K|.$$

By Property (14)

$$|S_K| = O(b_K q^N) = O(q^N q^{-\frac{N+1}{2}} \log^c q) = O(q^{\frac{N-1}{2}} \log^c q),$$

then

$$|\Omega_K| = O(q^{\frac{N-1}{2}} \log^c q).$$

Also, from Properties (3) and (11),

$$|B_i| = O(\theta q^{\frac{N-1}{2}}).$$

Since after the K -th step the uncovered points are a subset of $\Omega_K \cup (\cup_{i=1}^K B_i \setminus M_i)$, the size of this set is $O(q^{\frac{N-1}{2}} \log^c q)$ and therefore there exists a complete cap of size

$$O(q^{\frac{N-1}{2}} \log^c q).$$

□

6. Instruments

In this section we summarize all the instruments from [14] which will be used for the proofs of the properties of the first and the second phase in the Main Lemma.

One of the most important tool used in this paper is the polynomial method; see [14, Section 4.2]. In general, when using the probabilistic method, a key point is to show that a particular random variable is strongly concentrated around its mean. In [14] the authors pointed out that none of the existing (classical) tools was sufficiently strong to prove the fundamental properties in the nibble process; therefore they presented two concentration results (Lemmas 4.1 and 4.2 in [14]) we also use in this paper.

Let t_1, \dots, t_r be r independent binary random variables with expectation p_i . They indicate if the i -th point is chosen or not. Let Y be a function depending on t_1, \dots, t_r ; then for $i = 1, \dots, r$ let

$$C_i(v) = \left| \mathbb{E} \left(Y(v^{(1)}) - Y(v^{(0)}) | t_1, \dots, t_{i-1} \right) \right|,$$

where

$$Y(v^{(1)}) = Y(t_1, \dots, t_{i-1}, 1, t_{i+1}, \dots, t_r)$$

and

$$Y(v^{(0)}) = Y(t_1, \dots, t_{i-1}, 0, t_{i+1}, \dots, t_r).$$

$C_i(v)$ is called the conditional average effect of the random variable t_i when t_1, \dots, t_{i-1} are given. Let $L \subset S_i$ be a subset of size $|L| \geq \log^{100} q$ and $L' = L \cap S_{i+1}$. Given a point j , define

$$A(L, j) = \{t \in L \mid (tj) \cap A_i \neq \emptyset\} \quad \text{and} \quad a(L) = \max \left\{ \max_{j \in \Omega_i} |A(L, j)|, |L| \log^{-100} q \right\}.$$

In the following we list some of the results of [14] we will use in Sections 7 and 8. The proofs are all very similar to the ones in [14]. Also, in some cases, they are precisely the same (see e.g. Lemmas 6.4 and 6.1 and Corollary 6.5) and therefore they are omitted.

Lemma 6.1 (Lemma 4.2 in [14]). *Let H be a hypergraph with vertex $\mathcal{V}(H) = \{1, 2, \dots, r\}$ and edges $\mathcal{E}(H)$ such that each edge has at most k vertices. Let $w(e)$ the weight of a single edge e . Suppose t_1, \dots, t_r are independent random variables (binary random variables with expected value p_i). Consider*

$$Y_H = \sum_{e \in \mathcal{E}(H)} w(e) \prod_{s \in e} t_s.$$

Let

$$\mathbb{E}_i(Y) = \max_{A \subset \mathcal{V}(H) \mid |A|=i} \mathbb{E}(Y_{H_A}),$$

where Y_{H_A} is the A -truncated subgraph of H , that is the hypergraph with vertices $\mathcal{V}(H) \setminus A$, $\mathcal{E}(H_A) = \{B \subset \mathcal{V}(H_A), B \cup A \in \mathcal{E}(H)\}$ and $w(B) = w(B \cup A)$. Then there exist positive numbers c_k, d_k depending only on k so that for any $\lambda > 0$

$$\Pr \left(|Y - \mathbb{E}(Y)| \geq (EE')^{1/2} \lambda^k \right) \leq d_k e^{-\lambda + k \log n}.$$

Let $L \subset S_i$ (or $L \subset \Omega_i$). For any point j , let

$$A(L, j) = \{t \in L \mid (tj) \cap A \neq \emptyset\}.$$

Lemma 6.2 (Lemma 4.4 in [14]). *With very high probability,*

$$C_k(v) = o(\log q) a(L).$$

Proof. The proof is the same as in [14, Lemma 4.4], since $p_i |A_i(j)| \leq 2a_i \theta = o(1)$ and $2b_i q p_i \leq q^{-\frac{N-1}{2}} \leq \log^{-102} q$, for q large enough. \square

Lemma 6.3 (Lemma 4.5 in [14]). *With very high probability,*

$$\sum_{k=1}^{|S_i|} p_i C_k(v) = o(\log q) |L|.$$

Proof. The proof is the same as in [14, Lemma 4.5], since $p_i \max_j |A_i(j)| = o(1)$ and $2b_i q p_i \leq q^{-\frac{N-1}{2}} \leq \log^{-102} q$, for q large enough. \square

Lemma 6.4 (Lemma 4.7 in [14]). *Let $L \subset \Omega$. With very high probability,*

$$|L' - \mathbb{E}(L')| \leq a(L)^{1/2} |L|^{1/2} \log^5 q.$$

Corollary 6.5 (Corollary 4.8 in [14]). *Let K be a fixed positive constant. For any set L such that $a(L) \leq \frac{|L|}{\log^{2(K+5)} q}$, with very high probability,*

$$|L' - \mathbb{E}(L')| \leq |L| \log^{-K} q.$$

The following lemma states that, given two points $x, y \in S_i$, even if the events $x \in S_{i+1}$ and $y \in S_{i+1}$ are not independent, they are “almost independent”.

Lemma 6.6 (Lemma 5.3 in [14]). *Let $x, y \in S_i$, then*

$$|Pr(x, y \in S_{i+1}) - Pr(x \in S_{i+1})Pr(y \in S_{i+1})| = o(\log^{-13} q).$$

7. First Phase

Proposition 7.1. *Property (3) holds.*

Proof. By induction on i .

- $|B_i| \leq 2\theta q^{\frac{N-1}{2}}$: every point is chosen with the same probability $p_i = \frac{\theta}{b_i q^{\frac{N+1}{2}}}$ from the set S_i , then, since by induction $|S_i| \simeq (1 + \log^{-10} q)^i b_i q^N$ (see Property (18)), with very high probability

$$|B_i| \simeq p_i |S_i| = \frac{\theta}{b_i q^{\frac{N+1}{2}}} (1 + \log^{-10} q)^i b_i q^N = (1 + o(1)) \theta q^{\frac{N-1}{2}} \leq 2\theta q^{\frac{N-1}{2}}.$$

- $\theta q^{\frac{N-1}{2}} (1 - o(1)) \leq |M_i| \leq \theta q^{\frac{N-1}{2}} (1 + o(1))$: to show this, we get a bound on $U_i = B_i \setminus M_i$. Using the polynomial method (see [14, Subsection 4.2]), consider the variables

$$\{t_j | j \in S_i\},$$

such that $t_j = 1$ if $j \in B_i$. If a point is in B_i but not in M_i , then there exists a pair of points j', j'' collinear with j such that either $j', j'' \in B_i$ or $j' \in B_i$ and $j'' \in A_i$. In the first case

$$\sum_{j', j'' \in [j, j', j'']} t_{j'} t_{j''} > 0,$$

while in the second case

$$\sum_{j' \in A_i(j)} t_{j'} > 0.$$

We have

$$|U_i| = \sum_{j \in S_i} t_j \mathbf{1}_{j \notin M_i} \leq \sum_{j \in S_i} t_j \left(\sum_{j', j'' \in [j, j', j'']} t_{j'} t_{j''} + \sum_{j' \in A_i(j)} t_{j'} \right).$$

Consider

$$Y = \sum_{j \in S_i} t_j \sum_{j' \in A_i(j)} t_{j'} = \sum_{j \in S_i, j' \in A_i(j)} t_j t_{j'}.$$

To use Lemma 4.2 of [14], we have to compute $\mathbb{E}_i(Y)$, for $i = 0, 1, 2$. By the induction hypothesis $|S_i| \leq \sqrt{2}b_i q^N$ (Property (18)) and $|A_i(j)| \leq \sqrt{2}a_i b_i q^{\frac{N+1}{2}}$ (Property (17)). Then

$$\mathbb{E}_0(Y) = \mathbb{E}(Y) \leq \sqrt{2}a_i b_i q^{\frac{N+1}{2}} \sqrt{2}b_i q^N p_i^2 = 2a_i b_i^2 \frac{\theta^2}{b_i^2 q^{N+1}} q^N q^{\frac{N+1}{2}} = 2a_i \theta^2 q^{\frac{N-1}{2}}.$$

To bound $\mathbb{E}_1(Y)$, notice that for a fixed j there are at most $\max_j |A_i(j)|$ points j' such that $t_j t_{j'}$ appears in Y , so

$$\mathbb{E}_1(Y) \leq \sqrt{2}a_i b_i q^{\frac{N+1}{2}} p_i = \sqrt{2}a_i b_i q^{\frac{N+1}{2}} \frac{\theta}{b_i q^{\frac{N+1}{2}}} = \sqrt{2}a_i \theta \leq 2i\theta^2 = o(1),$$

since $a_i \simeq i\theta$. To bound $\mathbb{E}_2(Y)$, notice that each product $t_j t_{j'}$ can appear only at most twice in Y , so

$$\mathbb{E}_2(Y) \leq 2.$$

We can now apply Lemma 4.2 of [14] and we get that Y is strongly concentrate and in particular

$$Y \leq 3a_i \theta^2 q^{\frac{N-1}{2}} = o\left(\theta q^{\frac{N-1}{2}}\right)$$

with very high probability. Now consider

$$Y = \sum_{j \in S_i} t_j \sum_{j', j'' \in [j, j', j'']} t_{j'} t_{j''} = \sum_{j, j', j'' \in [j, j', j'']} t_j t_{j'} t_{j''}.$$

Arguing as before, we have

$$\begin{aligned} \mathbb{E}_0(Y) = \mathbb{E}(Y) &\leq |S_i|^2 |S_i((jj'))| p_i^3 \leq (\sqrt{2}b_i q^N)^2 \sqrt{2}b_i q p_i^3 = 2\sqrt{2}b_i^3 q^{2N+1} \frac{\theta^3}{b_i^3 q^{\frac{3N+3}{2}}} = 2\sqrt{2}\theta^3 q^{\frac{N-1}{2}} \\ \mathbb{E}_1(Y) &\leq |S_i| |S_i((jj'))| p_i^2 \leq \sqrt{2}b_i q^N \sqrt{2}b_i q p_i^2 = 2b_i^2 q^{N+1} \frac{\theta^2}{b_i^2 q^{N+1}} = 2\theta^2 \\ \mathbb{E}_2(Y) &\leq |S_i((jj'))| p_i \leq b_i q p_i = b_i q \frac{\theta}{b_i q^{\frac{N+1}{2}}} = o(1) \\ &\mathbb{E}_3(Y) \leq 6. \end{aligned}$$

Then, with very high probability

$$Y \leq 4\theta^3 q^{\frac{N-1}{2}} = o\left(\theta q^{\frac{N-1}{2}}\right),$$

which implies $|U_i| = o\left(\theta q^{\frac{N-1}{2}}\right)$ and therefore we have proven the bound on $|M_i|$.

□

Proposition 7.2. *Property (4) holds.*

Proof. We have to prove that

$$b_{i+1}q(1 - (i+1)\log^{-13}q) \leq |S_{i+1}(\ell)| \leq b_{i+1}q(1 + (i+1)\log^{-13}q).$$

Consider $L = S_i(\ell)$ and, in order to apply Corollary 6.5, $K = 14$. Let $x \in A(S_i(\ell), j)$ then: $x \in \ell$, $x \in S_i$, and there exists $a_1 \in A_i$ such that $[a_1 x j]$ (then $x \in S_i(\ell, j)$); so, by Property (6),

$$a(L) \leq \max_{\ell', v} |S_i(\ell', v)| = O(ia_i b_i \sqrt{q} + i \log^{40} q).$$

By the induction hypothesis, we have that $|L| = |S_i(\ell)| \geq \frac{1}{2}b_i q \geq \frac{1}{2} \log^{100} q$ (recall that we are in the first phase) then, since $a_i = o(\log^2 q)$,

$$a(L) \leq O(ia_i b_i \sqrt{q} + i \log^{40} q) \leq \frac{\frac{1}{2}b_i q}{\log^{38} q} \leq \frac{|L|}{\log^{2(K+5)} q}.$$

Therefore, by Corollary 6.5, with very high probability

$$||L'| - \mathbb{E}(|L'|)| = ||S_{i+1}(\ell)| - \mathbb{E}(|L'|)| \leq |L| \log^{-14} q \leq 2b_i q \log^{-14} q.$$

Note that (see also Remark 2.2) $\mathbb{E}(|L'|) = \frac{b_{i+1}}{b_i} |L|$ and then, by hypothesis,

$$|\mathbb{E}(|L'|) - b_{i+1}q| \leq ib_{i+1}q \log^{-13} q.$$

Finally we get from the two inequalities above

$$||L'| - b_{i+1}q| = ||S_{i+1}(\ell)| - b_{i+1}q| \leq ib_{i+1}q \log^{-13} q + 2b_i q \log^{-14} q \leq (i+1)b_{i+1}q \log^{-13} q.$$

□

Proposition 7.3. *Property (5) holds.*

Proof. The proof is similar to the proof of Proposition 7.2.

We have to prove that

$$|\Omega_{i+1}(\ell)| \leq b'_{i+1}q(1 + (i+1)\log^{-13}q).$$

Consider $L = \Omega_i(\ell)$ and, in order to apply Corollary 6.5, $K = 14$. Let $x \in A(\Omega_i(\ell), j)$ then: $x \in \ell$, $x \in \Omega_i$, and there exists $a_1 \in A_i$ such that $[a_1 x j]$ (then $x \in \Omega_i(\ell, j)$); so, by Property (6),

$$a(L) \leq \max_{\ell', v} |\Omega_i(\ell', v)| = O(ia_i b'_i \sqrt{q} + i \log^{40} q).$$

By the induction hypothesis, we have that $|L| = |\Omega_i(\ell)| \geq |S_i(\ell)| \geq \frac{1}{2} b_i q \geq \frac{1}{2} \log^{100} q$ (recall that we are in the first phase) then, since $a_i = o(\log^2 q)$,

$$a(L) \leq O(ia_i b'_i \sqrt{q} + i \log^{40} q) \leq \frac{\frac{1}{2} b_i q}{\log^{38} q} \leq \frac{|L|}{\log^{2(K+5)} q}.$$

Therefore, by Corollary 6.5, with very high probability

$$|L'| \leq \mathbb{E}(|L'|) + |L| \log^{-14} q \leq \mathbb{E}(|L'|) + 2b_i q \log^{-14} q.$$

Note that $\mathbb{E}(|L'|) \leq \frac{b'_{i+1}}{b'_i} |L|$ and then, by hypotesis,

$$|\mathbb{E}(|L'|) - b_{i+1} q| \leq i b'_{i+1} q \log^{-13} q.$$

Finally we get from the two inequalities above

$$|L'| = |\Omega_{i+1}(\ell)| \leq b'_{i+1} q + i b'_{i+1} q \log^{-13} q + 2b'_i q \log^{-14} q \leq b'_{i+1} q + (i+1) b'_{i+1} q \log^{-13} q.$$

□

Proposition 7.4. *Property (6) holds.*

Proof. We have to show that for every j

$$|S_j(\ell, v)| \leq 8j a_j b_j \sqrt{q} + j \log^{40} q.$$

Consider $L = S_i(\ell, v)$. By hypotesis

$$|S_i(\ell, v)| \leq 8i a_i b_i \sqrt{q} + i \log^{40} q.$$

Notice that if $x \in S_{i+1}(\ell, v)$ but not in $L' = L \cap S_{i+1}$, it means that in the line (xv) some point is chosen and belongs to A_{i+1} ; also,

$$|S_{i+1}(\ell, v)| = |L'| + |S_{i+1}(\ell, v) \setminus L'|.$$

Recall that

$$A(L, j) = \{t \in L \mid (tj) \cap A_i \neq \emptyset\}$$

and

$$a(L) = \max \left\{ \max_{j \in \Omega_i} |A(L, j)|, |L| \log^{-100} q \right\}.$$

Let $x \in A(S_i(\ell, v), j)$, then: $x \in \ell$, $x \in S_i$, there exists $a_1 \in A_i$ such that $[a_1 x v]$, and there exists $a_2 \in A_i$ such that $[a_2 x j]$. So, $x \in S_i(\ell, v, j)$. Therefore

$$a(L) \leq \max_{\ell, u, v} |S_i(\ell, v, u)| \leq i \log^4 q \leq \log^7 q.$$

Observe we used $i \leq \log^3 q$, which is true by hypothesis. Moreover (see Remark 2.2)

$$\mathbb{E}(|L'|) = |L|(1 - P_i^u) \leq 8ia_i b_{i+1} \sqrt{q} + i \log^{40} q,$$

since $b_{i+1} = b_i(1 - P_i^u)$. By Lemma 6.4, with very high probability

$$|L'| \leq \mathbb{E}(L') + a(L)^{1/2} |L|^{1/2} \log^5 q \leq 8ia_i b_{i+1} \sqrt{q} + i \log^{40} q + |L|^{1/2} \log^9 q.$$

For the other set, since in the line (xv) some point j has $t_j = 1$, we have

$$|S_{i+1}(\ell, v) \setminus L'| \leq \sum_{x \in S_i(\ell)} \sum_{j \in (xv)} t_j = Y.$$

As before,

$$\mathbb{E}_0(Y) = \mathbb{E}(Y) \leq p_i |S_i(\ell')|^2 \leq (\sqrt{2} b_i q)^2 p_i = 2b_i^2 q^2 \frac{\theta}{b_i q^{\frac{N+1}{2}}} = 2b_i \theta q^{-\frac{N-3}{2}} = o(1)$$

$$\mathbb{E}_1(Y) \leq 1.$$

(recall that $N \geq 3$). So, with very high probability

$$|S_{i+1}(\ell, v) \setminus L'| = O(\log^3 q).$$

Therefore

$$(8ia_i b_{i+1} \sqrt{q} + i \log^{40} q)^{1/2} \log^9 q \leq 8a_i b_{i+1} \sqrt{q},$$

so

$$(8ia_i b_{i+1} \sqrt{q} + i \log^{40} q)^{1/2} \log^9 q + \log^4 q \leq 8a_i b_{i+1} \sqrt{q} + \log^{40} q;$$

then

$$\begin{aligned} |S_{i+1}(\ell, v)| &\leq 8ia_i b_{i+1} \sqrt{q} + i \log^{40} q + 8a_i b_{i+1} \sqrt{q} + \log^{40} q \leq \\ &\leq 8(i+1)a_i b_{i+1} \sqrt{q} + (i+1) \log^{40} q \leq 8(i+1)a_{i+1} b_{i+1} \sqrt{q} + (i+1) \log^{40} q, \end{aligned}$$

since $a_{i+1} \simeq a_i + \theta$. \square

Proposition 7.5. *Property (7) holds.*

Proof. We have to prove that $|S_{i+1}(\ell, u, v)| \leq (i+1) \log^4 q$. Recall that a point x is in $S_{i+1}(\ell, u, v)$ if it lies on the line ℓ and there exist two points $a_1, a_2 \in A_{i+1}$ such that $[x, a_1, u]$ and $[x, a_2, v]$. Let $B_i(\ell, u)$ be the set of all the $x \in \ell$ such that there exists $b \in B_i$ with $[x, b, u]$ and $B_i(\ell, u, v) = B_i(\ell, u) \cap B_i(\ell, v)$. If $x \in S_{i+1}(\ell, u, v) \setminus S_i(\ell, u, v)$ it means that

$$x \in B_i(\ell, u, v) \cup (S_i(\ell, u) \cap B_i(\ell, v)) \cup (S_i(\ell, v) \cap B_i(\ell, u)),$$

so

$$|S_{i+1}(\ell, u, v)| \leq |S_i(\ell, u, v)| + |B_i(\ell, u, v)| + |S_i(\ell, u) \cap B_i(\ell, v)| + |S_i(\ell, v) \cap B_i(\ell, u)|.$$

Let $x \in B_i(\ell, u, v)$, then $x \in S_i(\ell)$ and there exist $j, j' \in B_i$ with $[x, u, j]$ and $[x, v, j']$ which are “chosen”, that is $t_j = t_{j'} = 1$. Then

$$|B_i(\ell, u, v)| \leq \sum_{x \in S_i(\ell)} \left(\sum_{j \in [xu, j]} t_j \right) \left(\sum_{j' \in [xv, j']} t_{j'} \right) = Y$$

and

$$\mathbb{E}_0(Y) = \mathbb{E}(Y) \leq |S_i(\ell')|^3 p_i^2 \leq (\sqrt{2}b_i q)^3 p_i^2 \leq 4b_i^3 q^3 \frac{\theta^2}{b_i^2 q^{N+1}} = 4b_i \theta^2 q^{-N+2} = o(1)$$

$$\mathbb{E}_1(Y) \leq |S_i(\ell')| p_i \leq \sqrt{2}b_i q p_i = \sqrt{2}b_i q \frac{\theta}{b_i q^{\frac{N+1}{2}}} = \sqrt{2}\theta q^{-\frac{N-1}{2}} = o(1)$$

$$\mathbb{E}_2(Y) \leq 2.$$

Thus, with very high probability (as noted after [14, Corollary 4.3]) $|B_i(\ell, u, v)| = O(\log^3 q)$.

For the second term,

$$|S_i(\ell, u) \cap B_i(\ell, v)| \leq \sum_{x \in S_i(\ell)} \left(\sum_{j \in A_i(u) \cap [xu, j]} t_j \right) \left(\sum_{j' \in [xv, j']} t_{j'} \right) = Y$$

and

$$\mathbb{E}_0(Y) = \mathbb{E}(Y) \leq |S_i(\ell')|^3 p_i^2 \leq 4b_i^3 q^3 p_i^2 = 4b_i^3 q^3 \frac{\theta^2}{b_i^2 q^{N+1}} = 4b_i \theta^2 q^{-N+2} = o(1)$$

$$\mathbb{E}_1(Y) \leq |S_i(\ell)| p_i = o(1)$$

$$\mathbb{E}_2(Y) \leq 2.$$

As above, $|S_i(\ell, u) \cap B_i(\ell, v)| = O(\log^3 q)$. The same for $|S_i(\ell, v) \cap B_i(\ell, u)|$. Since by induction $|S_i(\ell, u, v)| \leq i \log^4 q$, then

$$|S_{i+1}(\ell, u, v)| \leq (i+1) \log^4 q.$$

□

Proposition 7.6. *Property (8) holds.*

Proof. We have to prove that

$$|A_{i+1}(u, v)| \leq (i+1)b_{i+1}q + (i+1)\log^{40} q.$$

A point in $A_{i+1}(u, v)$ can be contained in $A_i(u, v)$, $B_i(u, v)$, $A_i(u) \cap B_i(v)$ or $A_i(v) \cap B_i(u)$. Then

$$|A_{i+1}(u, v)| \leq |A_i(u, v)| + |B_i(u, v)| + |A_i(u) \cap B_i(v)| + |A_i(v) \cap B_i(u)|.$$

- By the induction hypothesis $|A_i(u, v)| \leq ib_i q + i \log^{40} q$.
- $B_i(u, v)$: for any $x \in B_i(u, v)$ there exist $j, j' \in B_i$ such that $[xuj]$, $[xvj']$. So,

$$|B_i(u, v)| \leq \sum_{x \in S_i} \left(\sum_{j \in [xuj]} t_j \right) \left(\sum_{j' \in [xvj']} t_{j'} \right) = Y.$$

In order to evaluate $\mathbb{E}_i(Y)$, notice that if j is fixed, then $x \in S_i((ju))$, while j' belongs to $S_i((xv))$. Then

$$\mathbb{E}_0(Y) = \mathbb{E}(Y) \leq p_i^2 |S_i| (\max_{\ell} |S_i(\ell)|)^2 = \frac{\theta^2}{b_i^2 q^{N+1}} 2b_i q^N (2b_i q)^2 = 8b_i \theta^2 q$$

$$\mathbb{E}_1(Y) \leq p_i (\max_{\ell} |S_i(\ell)|)^2 = \frac{\theta}{b_i q^{\frac{N+1}{2}}} (2b_i q)^2 = 4b_i \theta q^{-\frac{N-3}{2}} = o(1)$$

$$\mathbb{E}_2(Y) \leq 2.$$

By Lemma 6.1 and using $\lambda = \log^{3/2} q$ we get $|B_i(X)| \leq 8b_i \theta^2 q + 4\theta \sqrt{q} \log^3 q = o(b_i q \log^{-3} q)$.

- $A_i(u) \cap B_i(v)$: for any $x \in A_i(u) \cap B_i(v)$ there exist $j_1 \in A_i$ and $j_2 \in B_i$ such that $[xuj_1]$ and $[xvj_2]$.

$$|A_i(u) \cap B_i(v)| \leq \sum_{x \in (uj_1)} \left(\sum_{j_2 \in [xvj_2]} t_{j_2} \right) = Y.$$

Then

$$\mathbb{E}_0(Y) \leq p_i |A_i(u)| (\max_{\ell} |S_i(\ell)|) = \frac{\theta}{b_i q^{\frac{N+1}{2}}} 2ia_i b_i q^{\frac{N+1}{2}} (2b_i q) = 4a_i b_i \theta q$$

$$\mathbb{E}_1(Y) = 1.$$

By Lemma 6.1 and using $\lambda = \log^{1/2} q$ we get $|A_i(u) \cap B_i(v)| \leq 4b_i \theta^2 q + 2i\theta \sqrt{q} = o(b_i q \log^{-3} q)$. Therefore

$$|A_{i+1}(u, v)| \leq ib_i q + i \log^{40} q + o(b_i q \log^{-3} q) \leq (i+1)b_{i+1} q + (i+1) \log^{40} q.$$

□

Proposition 7.7. *Property (9) holds.*

Proof. The proof is very similar to the proof of Proposition 7.4.

We have to show that for every j

$$|\Omega_j(\ell, v)| \leq 8ja_j b'_j \sqrt{q} + j \log^{40} q.$$

Consider $L = \Omega_i(\ell, v)$. By hypothesis $|\Omega_i(\ell, v)| \leq 8ia_i b'_i \sqrt{q} + i \log^{40} q$. Notice that if $x \in \Omega_{i+1}(\ell, v)$ but not in $L' = L \cap \Omega_{i+1}$, it means that in the line (xv) some point is chosen and belongs to A_{i+1} ; also,

$$|\Omega_{i+1}(\ell, v)| = |L'| + |\Omega_{i+1}(\ell, v) \setminus L'|.$$

Recall that

$$A(L, j) = \{t \in L \mid (tj) \cap A_i \neq \emptyset\}$$

and

$$a(L) = \max \left\{ \max_{j \in \Omega_i} |A(L, j)|, |L| \log^{-100} q \right\}.$$

Let $x \in A(\Omega_i(\ell, v), j)$, then: $x \in \ell$, $x \in \Omega_i$, there exists $a_1 \in A_i$ such that $[a_1 xv]$, and there exists $a_2 \in A_i$ such that $[a_2 xj]$. So, $x \in \Omega_i(\ell, v, j)$. Therefore

$$a(L) \leq \max_{\ell, u, v} |\Omega_i(\ell, v, u)| \leq i \log^4 q \leq \log^7 q.$$

Notice that we used that $i \leq \log^3 q$ which is true by hypothesis. Moreover

$$\mathbb{E}(|L'|) \leq |L|(1 - P_i^\ell) \leq 8ia_i b'_{i+1} \sqrt{q} + i \log^{40} q,$$

since $b'_{i+1} = b'_i(1 - P_i^\ell)$. By Lemma 6.4, with very high probability

$$|L'| \leq \mathbb{E}(L') + a(L)^{1/2} |L|^{1/2} \log^5 q \leq 8ia_i b'_{i+1} \sqrt{q} + i \log^{40} q + |L|^{1/2} \log^9 q.$$

For the other set, since in the line (xv) some point j has $t_j = 1$, we have

$$|\Omega_{i+1}(\ell, v) \setminus L'| \leq \sum_{x \in \ell} \sum_{j \in (xv)} t_j = Y.$$

As before

$$\begin{aligned} \mathbb{E}_0(Y) = \mathbb{E}(Y) &\leq p_i |\Omega_i(\ell')|^2 \leq (\sqrt{2} b'_i q)^2 p_i = 2(b'_i)^2 q^2 \frac{\theta}{b_i q^{\frac{N+1}{2}}} \leq 4b_i \theta q^{-\frac{N-3}{2}} = o(1) \\ \mathbb{E}_1(Y) &\leq 1. \end{aligned}$$

(recall that $N \geq 3$). So, with very high probability

$$|\Omega_{i+1}(\ell, v) \setminus L'| = O(\log^3 q).$$

Therefore

$$(8ia_i b'_i \sqrt{q} + i \log^{40} q)^{1/2} \log^9 q \leq 8a_i b'_{i+1} \sqrt{q},$$

so

$$(8ia_i b'_i \sqrt{q} + i \log^{40} q)^{1/2} \log^9 q + \log^{40} q \leq 8a_i b'_{i+1} \sqrt{q} + \log^{40} q;$$

then

$$\begin{aligned} |\Omega_{i+1}(\ell, v)| &\leq 8ia_i b'_{i+1} \sqrt{q} + i \log^{40} q + 8a_i b'_{i+1} \sqrt{q} + \log^{40} q = \\ &= 8(i+1)a_i b'_{i+1} \sqrt{q} + (i+1) \log^{40} q \leq 8(i+1)a_{i+1} b'_{i+1} \sqrt{q} + (i+1) \log^{40} q, \end{aligned}$$

since $a_{i+1} \simeq a_i + \theta$. \square

Proposition 7.8. *Property (10) holds.*

Proof. The proof is similar to the proof of Proposition 7.5.

We have to prove that $|\Omega_{i+1}(\ell, u, v)| \leq (i+1) \log^4 q$. Recall that a point $x \in \Omega_i$ is in $\Omega_{i+1}(\ell, u, v)$ if it lies on the line ℓ and there exist two points $a_1, a_2 \in A_{i+1}$ such that $[x, a_1, u]$ and $[x, a_2, v]$. Let $B_i(\ell, u)$ be the set of all the $x \in \ell$ such that there exists $b \in B_i$ with $[x, b, u]$ and $B_i(\ell, u, v) = B_i(\ell, u) \cap B_i(\ell, v)$. If $x \in \Omega_{i+1}(\ell, u, v) \setminus \Omega_i(\ell, u, v)$ it means that

$$x \in B_i(\ell, u, v) \cup (\Omega_i(\ell, u) \cap B_i(\ell, v)) \cup (\Omega_i(\ell, v) \cap B_i(\ell, u)),$$

so

$$|\Omega_{i+1}(\ell, u, v)| \leq |\Omega_i(\ell, u, v)| + |B_i(\ell, u, v)| + |\Omega_i(\ell, u) \cap B_i(\ell, v)| + |\Omega_i(\ell, v) \cap B_i(\ell, u)|.$$

Let $x \in B_i(\ell, u, v)$, then there exist $j, j' \in B_i$ with $[x, u, j]$ and $[x, v, j']$ which are “chosen”, that is $t_j = t_{j'} = 1$. Then

$$|B_i(\ell, u, v)| \leq \sum_{x \in \ell} \left(\sum_{j \in [xu]} t_j \right) \left(\sum_{j' \in [xv]} t_{j'} \right) = Y.$$

We already know from the proof of the Main Theorem, that

$$\log \frac{b'_i}{b_i} = \log \prod_{j=0}^i \frac{1 - P_j^\ell}{1 - P_j^u} = O(i \log^{-3} q) = o(1) \implies b'_i \leq 2b_i.$$

Then

$$\begin{aligned} \mathbb{E}_0(Y) &= \mathbb{E}(Y) \leq |\Omega_i(\ell')|^3 p_i^2 \leq (\sqrt{2}b'_i q)^3 p_i^2 \leq 4(b'_i)^3 q^3 \frac{\theta^2}{b_i^2 q^{N+1}} \leq 32b_i \theta^2 q^{-N+2} = o(1) \\ \mathbb{E}_1(Y) &\leq |\Omega_i(\ell')| p_i \leq \sqrt{2}b'_i q p_i = \sqrt{2}b'_i q \frac{\theta}{b_i q^{\frac{N+1}{2}}} = 2\sqrt{2}\theta q^{-\frac{N-1}{2}} = o(1) \\ \mathbb{E}_2(Y) &\leq 2. \end{aligned}$$

With very high probability (as noted after [14, Corollary 4.3]) $|B_i(\ell, u, v)| = O(\log^3 q)$. Also, consider

$$|\Omega_i(\ell, u) \cap B_i(\ell, v)| \leq \sum_{x \in \ell} \left(\sum_{j \in A_i(u) \cap [xu]} t_j \right) \left(\sum_{j' \in [xv]} t_{j'} \right) = Y.$$

Then

$$\begin{aligned} \mathbb{E}_0(Y) &= \mathbb{E}(Y) \leq |\Omega_i(\ell')|^3 p_i^2 \leq (\sqrt{2}b'_i q)^3 p_i^2 \leq 4(b'_i)^3 q^3 \frac{\theta^2}{b_i^2 q^{N+1}} \leq 32b_i \theta^2 q^{-N+2} = o(1) \\ \mathbb{E}_1(Y) &\leq |\Omega_i(\ell)| p_i = o(1) \\ \mathbb{E}_2(Y) &\leq 2. \end{aligned}$$

As above, $|\Omega_i(\ell, u) \cap B_i(\ell, v)| = O(\log^3 q)$. The same for $|\Omega_i(\ell, v) \cap B_i(\ell, u)|$. Since by induction $|\Omega_i(\ell, u, v)| \leq i \log^4 q$, thus

$$|\Omega_{i+1}(\ell, u, v)| \leq (i+1) \log^4 q.$$

□

8. Second phase

Remark 8.1. *Property 7.6 is still valid in Phase two, since in the proof we used properties which are still valid in this phase. Also, let \bar{i} such that $b_{\bar{i}+1} < \frac{\log^{c_1} q}{q} \leq b_{\bar{i}}$. Then for $i > \bar{i} + 1$*

$$|S_i(\ell)| \leq |S_{\bar{i}+1}(\ell)| \leq 2b_{\bar{i}+1}q \leq 2 \log^{c_1} q$$

and

$$|\Omega_i(\ell)| \leq |\Omega_{\bar{i}+1}(\ell)| \leq 2b'_{\bar{i}+1}q \leq 4b_{\bar{i}+1}q \leq \log^{c_1} q.$$

Proposition 8.2. *Property (11) holds.*

Proof. The proof is the same of Proposition 7.1. □

Proposition 8.3. *Property (12) holds.*

Proof. We have to prove that with very high probability

$$\frac{1}{2}b_{i+1}^2q^{N+1}(1 - 3(i+1)\log^{-13}q) \leq |T_{i+1}(v)| \leq \frac{1}{2}b_{i+1}^2q^{N+1}(1 + 3(i+1)\log^{-13}q).$$

Recall that $T_i(v) = \{\{x, y\} | x, y \in S_i \setminus \{v\}, [xyv]\}$, then

$$\mathbb{E}(T_{i+1}(v)) = \sum_{x, y \in S_i, [xyv]} Pr(x, y \in S_{i+1}).$$

Therefore, by Lemma 6.6,

$$\begin{aligned} |T_i(v)|Pr(x \in S_{i+1} | x \in S_i)^2(1 - o(\log^{-13}q)) &\leq \mathbb{E}(T_{i+1}(v)) \leq \\ &\leq |T_i(v)|Pr(x \in S_{i+1} | x \in S_i)^2(1 + o(\log^{-13}q)). \end{aligned}$$

By Remark 2.2 $Pr(x \in S_{i+1} | x \in S_i) = 1 - P_i^u$, so $b_i Pr(x \in S_{i+1} | x \in S_i) = b_{i+1}$, and by the induction hypothesis

$$\begin{aligned} \frac{1}{2}q^{N+1}b_{i+1}^2(1 - 3i \log^{-13}q)(1 - o(\log^{-13}q)) &\leq \mathbb{E}(T_{i+1}(v)) \leq \\ &\leq \frac{1}{2}q^{N+1}b_{i+1}^2(1 + 3i \log^{-13}q)(1 + o(\log^{-13}q)). \end{aligned}$$

Consider

$$T'_i(v) = \{x^{m(x)} | x \in S_i \setminus \{v\}\},$$

where $m(x) = |S_i((x, v))| - 2$. It is easy to see that $|T'_i(v)| = 2|T_i(v)|$. Let $L = T'_i(v)$ and $L' = T'_{i+1}(v)$, it holds that

$$|A(L, u)| = |\{t \in L | (tu) \cap A \neq \emptyset\}| \leq |A_i(u)| \max_x m(x),$$

since every element $x \in A_i(u)$ is repeated in $A(L, u)$ exactly $m(x)$ times. Moreover, since we are in the second phase and $b_i q \leq \log^{c_1} q$,

$$|A_i(u)| \max_x m(x) \leq |A_i(u)| \max_\ell |S_i(\ell)| \leq \sqrt{2} a_i b_i q^{\frac{N+1}{2}} \sqrt{2} \log^{c_1} q = 2a_i b_i q^{\frac{N+1}{2}} \log^{c_1} q.$$

Since $|L| = |T'_i(v)| \geq \frac{1}{3} b_i^2 q^{N+1} \geq \frac{a(L)}{\log^{100} q}$, by Lemma 6.4 we have that with very high probability

$$\begin{aligned} ||L'| - \mathbb{E}(|L'|)| &\leq (|L| a(L))^{1/2} \log^5 q \leq \left(2b_i^2 q^{N+1} 2a_i b_i q^{\frac{N+1}{2}} \log^{c_1} q \right)^{1/2} \log^5 q = \\ &= 2a_i^{1/2} b_i^{3/2} q^{\frac{3(n+1)}{4}} \log^{55} q \leq b_i^{3/2} q^{\frac{3(n+1)}{4}} \log^{60} q. \end{aligned}$$

Since $b_i q^{\frac{N+1}{2}} \geq \log^c q$ and $2b_{i+1} \geq b_i$, we have that

$$\frac{b_i^{3/2} q^{\frac{3(n+1)}{4}} \log^{60} q}{b_{i+1}^2 q^{N+1} \log^{-13} q} \leq 4 \frac{b_i^{3/2} q^{\frac{3(n+1)}{4}} \log^{60} q}{b_i^2 q^{N+1} \log^{-13} q} \leq 4 \left(b_i q^{\frac{N+1}{2}} \right)^{-1/2} \log^{73} q \leq 4 \log^{-77} q.$$

So $b_i^{3/2} q^{\frac{3(n+1)}{4}} \log^{60} q = o(b_{i+1}^2 q^{N+1} \log^{-13} q)$. Therefore with very high probability

$$\begin{aligned} b_{i+1}^2 q^{N+1} (1 - 3i \log^{-13} q) (1 - o(\log^{-13} q)) - o(b_{i+1}^2 q^{N+1} \log^{-13} q) &\leq |L'| \leq \\ &\leq b_{i+1}^2 q^{N+1} (1 + 3i \log^{-13} q) (1 + o(\log^{-13} q)) + o(b_{i+1}^2 q^{N+1} \log^{-13} q). \end{aligned}$$

We have that

$$\begin{aligned} b_{i+1}^2 q^{N+1} (1 - 3i \log^{-13} q) (1 - o(\log^{-13} q)) - o(b_{i+1}^2 q^{N+1} \log^{-13} q) &= \\ = b_{i+1}^2 q^{N+1} (1 - 3i \log^{-13} q - 3o(\log^{-13} q)) &\geq b_{i+1}^2 q^{N+1} (1 - 3(i+1) \log^{-13} q) \end{aligned}$$

and

$$\begin{aligned} b_{i+1}^2 q^{N+1} (1 + 3i \log^{-13} q) (1 + o(\log^{-13} q)) + o(b_{i+1}^2 q^{N+1} \log^{-13} q) &= \\ = b_{i+1}^2 q^{N+1} (1 + 3i \log^{-13} q + 3(\log^{-13} q)) &\leq b_{i+1}^2 q^{N+1} (1 + 3(i+1) \log^{-13} q). \end{aligned}$$

Recalling that $|L'| = 2|T_i(v)|$ we have the result. \square

Proposition 8.4. *Property (13) holds.*

Proof. We have to prove that

$$a_{i+1}b_{i+1}q^{\frac{N+1}{2}}(1 - O(i\theta^2)) \leq |A_{i+1}(v)| \leq a_{i+1}b_{i+1}q^{\frac{N+1}{2}}(1 + O(i\theta^2)).$$

Let $U_i = B_i \setminus M_i$ and $U_i(v), B_i(v), M_i(v)$ the set of points $x \in S_i$ such that there exists $u \in U_i, B_i, M_i$ respectively with $[xuv]$, and $B'_i(v) = B_i(v) \cap S_{i+1}$, $M'_i(v) = M_i(v) \cap S_{i+1}$. Let $L = A_i(v)$, $L' = L \cap S_{i+1}$, $v \in \Omega_{i+1}$, then

$$|A_{i+1}(v)| = |L'| + |M'_i(v)|.$$

In the following we will give some estimations for the sizes of $B_i(v)$, $U_i(v)$, and $B'_i(v)$.

- $|B_i(v)| \leq \theta b_i q^{\frac{N+1}{2}}(1 + o(\log^{-10} q))$.
If $x \in B_i(v)$, then there exists $j \in B_i$ such that $[jxv]$ and $t_j = 1$. Then

$$|B_i(v)| \leq \sum_{x \in S_i} \left(\sum_{j \in S_i | [jxv]} t_j \right) = \sum_{j \in S_i} t_j m(j) = Y,$$

where $m(j) = |S_i((vj)) \setminus \{v, j\}|$. Notice that $\sum_{j \in S_i} m(j) = |T'_i(v)| = 2|T_i(v)|$. By Property (12),

$$\begin{aligned} \mathbb{E}_0(Y) = \mathbb{E}(Y) &= 2p_i |T_i(v)| \geq p_i b_i^2 q^{N+1} (i - o(\log^{-10} q)) \simeq \theta b_i q^{\frac{N+1}{2}} \geq \log^{150} q \\ \mathbb{E}_1(Y) &\leq \max_x m(x) \leq \max_{\ell} |S_i(\ell)| \leq 2 \log^{c_1} q \end{aligned}$$

By Lemma 6.1, with very high probability,

$$Y = |B_i(v)| \leq \mathbb{E}(Y) + (2\mathbb{E}(Y) \log^{c_1} q)^{1/2} \log^2 q \leq \theta b_i q^{\frac{N+1}{2}} (1 + o(\log^{-10} q)).$$

- $|U_i(v)| \leq 6\theta^2 a_i b_i q^{\frac{N+1}{2}}$.
Note that for any $x \in U_i(v)$ there exists $j \in (xv)$ different from x, v such that $j \in B_i \setminus M_i$; so

$$|U_i(v)| \leq \sum_{x \in S_i | |(xv)| \geq 3} \sum_{j \in S_i | [jxv]} t_j \mathbf{1}_{\{j \notin M_i\}}.$$

Moreover, from the description of the algorithm $j \notin M_i$ if and only if $[jj', j'']$ for either some $j', j'' \in B_i$ or $j' \in B_i$ and $j'' \in A_i$. Then

$$\sum_{x \in S_i | |(xv)| \geq 3} \sum_{j \in S_i | [jxv]} t_j \mathbf{1}_{\{j \notin M_i\}} \leq \sum_{x \in S_i | |(xv)| \geq 3} \sum_{j \in S_i | [jxv]} t_j \left(\sum_{j', j'' | [jj'j'']} t_{j'} t_{j''} + \sum_{j' \in A_i(j)} t_{j'} \right).$$

Consider

$$Y_1 = \sum_{x \in S_i | |(xv)| \geq 3} \sum_{j \in S_i | [jxv]} t_j \sum_{j', j'' | [jj'j'']} t_{j'} t_{j''}$$

and

$$Y_2 = \sum_{x \in S_i | |(xv)| \geq 3} \sum_{j \in S_i | [jxv]} t_j \sum_{j' \in A_i(j)} t_{j'}. \quad (22)$$

We have that

$$\begin{aligned} \mathbb{E}_0(Y_1) &\simeq p_i^3 |S_i| \max_{\ell} |S_i(\ell)| \max_j |T_i(j)| \simeq \left(\frac{\theta^3}{b_i^3 q^{\frac{3N+3}{2}}} \right) b_i^4 q^{2N+2} \simeq \theta^3 b_i q^{\frac{N+1}{2}} \\ \mathbb{E}_0(Y_2) &\simeq p_i^2 |S_i| \max_x |A_i(x)| \max_{\ell} |S_i(\ell)| \simeq \left(\frac{\theta^2}{b_i^2 q^{N+1}} \right) a_i b_i^3 q^{\frac{3N+3}{2}} \simeq a_i b_i \theta^2 q^{\frac{N+1}{2}}. \end{aligned}$$

It is easy to prove that with very high probability

$$Y_1 \leq 5\theta^3 b_i q^{\frac{N+1}{2}}, \quad Y_2 \leq 5\theta^2 a_i b_i q^{\frac{N+1}{2}};$$

therefore

$$|U_i(v)| \leq 6\theta^2 a_i b_i q^{\frac{N+1}{2}}.$$

- $|B'_i(v)| \geq \theta b_i q^{\frac{N+1}{2}} (1 - o(\log^{-10} q)) - 14a_i \theta^2 b_i q^{\frac{N+1}{2}}.$
A point $x \in B'_i(v)$ if and only if $x \in S_{i+1}$ and there exists at least one point in the line (xv) different from x and v belonging to B_i . Let $B''_i(v) = \{x \in S_{i+1} | \exists ! j \in B_i, j \neq x, v[jxv]\}$. Then

$$|B'_i(v)| \geq \sum_{x \in S_i | |(xv)| \geq 3} \mathbf{1}_{\{x \in S_{i+1}\}} \sum_{j | [jxv]} t_j \left(1 - \sum_{j' \neq j | [j'xv]} t_{j'} \right)$$

and

$$|B'_i(v)| \geq -|U_i(v)| + \sum_{x \in S_i | |(xv)| \geq 2} \left(1 - \left(\sum_{g \in A_i(x)} t_g + \sum_{g, g' | [gg'x]} t_g t_{g'} \right) \right) \sum_{j | [jxv]} t_j \left(1 - \sum_{j' \neq j | [j'xv]} t_{j'} \right).$$

In fact $\sum_{j | [jxv]} t_j \left(1 - \sum_{j' \neq j | [j'xv]} t_{j'} \right)$ is different from zero only if $|(xv)| = 3$; in this case

$$\left(1 - \left(\sum_{g \in A_i(x)} t_g + \sum_{g, g' | [gg'x]} t_g t_{g'} \right) \right)$$

is different from zero if and only if $x \in M_i(v) \cap U_i(v)$.

Consider

$$Y = \sum_{x \in S_i | |(xv)| \geq 2} \sum_{j | [jxv]} t_j.$$

Using the same calculations as above, we get that with very high probability

$$|Y| \geq \mathbb{E}_0(Y) (1 - o(\log^{-10} q)) \geq \theta b_i q^{\frac{N+1}{2}} (1 - o(\log^{-10} q)) (1 - o(\log^{-10} q)) \geq$$

$$\geq \theta b_i q^{\frac{N+1}{2}} (1 - o(\log^{-10} q)).$$

Consider the following

$$\begin{aligned} Y_1 &= - \sum_{x \in S_i | |(xv)| \geq 2} \sum_{j \in [jxv]} \sum_{j' \neq j | [j'xv]} t_j t_{j'}, \\ Y_2 &= - \sum_{x \in S_i | |(xv)| \geq 2} \sum_{g \in A_i(x)} \sum_{j \in S_i | [jxv]} t_g t_j, \\ Y_3 &= \sum_{x \in S_i | |(xv)| \geq 2} \sum_{g \in A_i(x)} \sum_{j \in S_i | [jxv]} \sum_{j' \neq j | [j'xv]} t_g t_j t_{j'}, \\ Y_4 &= - \sum_{x \in S_i | |(xv)| \geq 2} \sum_{g, g' | [gg'x]} \sum_{j \in S_i | [jxv]} t_g t_{g'} t_j, \\ Y_5 &= \sum_{x \in S_i | |(xv)| \geq 2} \sum_{g, g' | [gg'x]} \sum_{j \in S_i | [jxv]} \sum_{j' \neq j | [j'xv]} t_g t_{g'} t_j t_{j'}. \end{aligned}$$

We have that

$$\begin{aligned} \mathbb{E}_0(Y_1) &\simeq p_i^2 |S_i| \max_{\ell} |S_i(\ell)| \simeq \left(\frac{\theta^2}{b_i^2 q^{N+1}} \right) b_i^3 q^{N+2} \simeq \theta^2 b_i q, \\ \mathbb{E}_0(Y_2) &\simeq p_i^2 |S_i| \max_x |A_i(x)| \max_{\ell} |S_i(\ell)| \simeq \left(\frac{\theta^2}{b_i^2 q^{N+1}} \right) a_i b_i^3 q^{\frac{3n+3}{2}} \simeq a_i b_i \theta^2 q^{\frac{N+1}{2}}, \\ \mathbb{E}_0(Y_3) &\simeq p_i^3 |S_i| \max_x |A_i(x)| \left(\max_{\ell} |S_i(\ell)| \right)^2 \simeq \left(\frac{\theta^3}{b_i^3 q^{\frac{3n+3}{2}}} \right) a_i b_i^4 q^{\frac{3n+5}{2}} \simeq a_i b_i \theta^3 q, \\ \mathbb{E}_0(Y_4) &\simeq p_i^3 |S_i| \left(\max_{\ell} |S_i(\ell)| \right)^3 \simeq \left(\frac{\theta^3}{b_i^3 q^{\frac{3n+3}{2}}} \right) b_i^4 q^{N+3} = o(1), \\ \mathbb{E}_0(Y_5) &\simeq p_i^4 |S_i| \left(\max_{\ell} |S_i(\ell)| \right)^4 \simeq \left(\frac{\theta^4}{b_i^4 q^{2n+2}} \right) b_i^4 q^{N+4} = o(1). \end{aligned}$$

Then it is easy to see that the maximum of the expectations is $\mathbb{E}_0(Y_2)$ and the maximum of the errors is $\simeq a_i b_i \theta^2 q^{\frac{N+1}{2}}$. So

$$\begin{aligned} |B'_i(v)| &\geq \theta b_i q^{\frac{N+1}{2}} (1 - o(\log^{-10} q)) - 8a_i \theta^2 b_i q^{\frac{N+1}{2}} - 6a_i \theta^2 b_i q^{\frac{N+1}{2}} = \\ &= \theta b_i q^{\frac{N+1}{2}} (1 - o(\log^{-10} q)) - 14a_i \theta^2 b_i q^{\frac{N+1}{2}}. \end{aligned}$$

Note that since

$$1 - P_i^u = 1 - p_i \max_v |A_i(v)| (1 + o(1)) \geq 1 - \frac{\theta}{b_i q^{\frac{N+1}{2}}} \frac{3}{2} a_i b_i q^{\frac{N+1}{2}} \geq 1 - \frac{3}{2} \theta a_i,$$

it holds that

$$\frac{b_i}{b_{i+1}} = \frac{1}{1 - P_i^u} \leq \frac{1}{1 - \frac{3}{2} \theta a_i} \leq 1 + 2\theta a_i.$$

First of all observe that

$$|B'_i(v)| - |U_i(v)| \leq |M'_i(v)| \leq |B'_i(v)| \leq |B_i(v)|.$$

Then

$$\begin{aligned} |M'_i(v)| &\leq |B_i(v)| \leq \theta b_i q^{\frac{N+1}{2}} (1 + o(\log^{-10} q)) \leq \\ &\leq \theta b_{i+1} q^{\frac{N+1}{2}} (1 + 2\theta a_i) (1 + o(\log^{-10} q)) \leq \theta b_{i+1} q^{\frac{N+1}{2}} (1 + 3\theta a_i). \end{aligned}$$

Also,

$$\begin{aligned} |M'_i(v)| &\geq |B'_i(v)| - |U_i(v)| \geq \theta b_i q^{\frac{N+1}{2}} (1 - o(\log^{-10} q)) - 14a_i \theta^2 b_i q^{\frac{N+1}{2}} - 6\theta^2 a_i b_i q^{\frac{N+1}{2}} \geq \\ &\geq \theta b_i q^{\frac{N+1}{2}} (1 - 15a_i \theta) \geq \theta b_{i+1} q^{\frac{N+1}{2}} (1 - 21a_i \theta). \end{aligned}$$

Finally we have to estimate $|L'|$. Recall that $L = A_i(v)$, so $A(L, j) = A_i(v, j)$ and $a(L) = \max_{u,v} |A_i(u, v)|$. By Remark 8.1 we have that, since we are in the second phase,

$$a(L) \leq i b_i q + i \log^{40} q \leq \frac{1}{2} \log^3 q \log^{100} q + \log^3 q + \log^{40} q \leq \log^{103} q.$$

Also,

$$\mathbb{E}(|L'|) \simeq |A_i(v)| \frac{b_{i+1}}{b_i} \simeq a_i b_{i+1} q^{\frac{N+1}{2}} \geq a_i \log^c q \geq \log^{298} q.$$

Note that for some $Z > 0$, $a_i b_{i+1} q^{\frac{N+1}{2}} (1 - Z(i-1)\theta^2) \leq \mathbb{E}(|L'|) \leq a_i b_{i+1} q^{\frac{N+1}{2}} (1 + Z(i-1)\theta^2)$ by induction. Since $a(L) \leq \log^{103} q \leq \frac{\log^{c-2} q}{2 \log^{2(K+5)} q} \leq \frac{a_i b_i q^{\frac{N+1}{2}}}{2 \log^{2(K+5)} q} \leq \frac{|L|}{\log^{2(K+5)} q}$, by Corollary 6.5, with very high probability

$$a_i b_{i+1} q^{\frac{N+1}{2}} (1 - Z(i-1)\theta^2) (1 - o(\log^{-10} q)) \leq |L'| \leq a_i b_{i+1} q^{\frac{N+1}{2}} (1 + Z(i-1)\theta^2) (1 + o(\log^{-10} q)).$$

The lower bound on $|A_{i+1}(v)|$ is then

$$\begin{aligned} |A_{i+1}(v)| &= |L'| + |M'_i(v)| \geq a_i b_{i+1} q^{\frac{N+1}{2}} (1 - Z(i-1)\theta^2) (1 - o(\log^{-10} q)) + \theta b_{i+1} q^{\frac{N+1}{2}} (1 - 21a_i \theta) \geq \\ &\geq b_{i+1} q^{\frac{N+1}{2}} (a_i + \theta - Z a_i i \theta^2 - 21a_i \theta^2) \geq b_{i+1} q^{\frac{N+1}{2}} (a_{i+1} - \theta^3 - Z a_i i \theta^2 - 21a_i \theta^2), \end{aligned}$$

see also [14, Remark 5.1]; since $a_{i+1} \geq a_i$ and we can consider $Z \geq 22$ we have

$$|A_{i+1}(v)| \geq a_{i+1} b_{i+1} q^{\frac{N+1}{2}} (1 - Z(i+1)\theta^2).$$

Also, the upper bound on $|A_{i+1}(v)|$ is

$$\begin{aligned} |A_{i+1}(v)| &= |L'| + |M'_i(v)| \leq a_i b_{i+1} q^{\frac{N+1}{2}} (1 + Z(1-i)\theta^2) (1 + o(\log^{-10} q)) + \theta b_{i+1} q^{\frac{N+1}{2}} (1 + 3a_i \theta) \leq \\ &b_{i+1} q^{\frac{N+1}{2}} (a_i + \theta + Z a_i i \theta^2 + 3a_i \theta^2) \leq a_{i+1} b_{i+1} q^{\frac{N+1}{2}} (1 + Z i \theta^2 + 7\theta) \leq a_{i+1} b_{i+1} q^{\frac{N+1}{2}} (1 + Z(i+1)\theta^2), \end{aligned}$$

since, by [14, Remark 5.1] $a_{i+1} \geq a_i + \theta - (3a_i \theta^2 + 10\theta^3) \geq a_i + \theta - 4a_i \theta^2$. \square

Proposition 8.5. *Property (14) holds.*

Proof. We have to prove that with very high probability

$$(1 - \log^{-10} q)^{i+1} q^N b_{i+1} \leq |S_{i+1}| \leq (1 + \log^{-10} q)^{i+1} q^N b_{i+1}.$$

By Remark 8.1,

$$|S_i(\ell)| \leq 2 \log^{c_1} q.$$

Let $L = S_i$ and $L' = S_{i+1}$, $A(L, j) = \{x \in S_i \mid (jx) \cap A_i \neq \emptyset\} = A_i(j)$. An upper bound for $A(L, j)$ is, by Property (13),

$$a(L) \leq 2a_i b_i q^{\frac{N+1}{2}} \leq 2a_i q^{\frac{N-1}{2}} \log^{c_1} q.$$

By hypothesis $|L| \geq \frac{1}{2} b_i q^N \geq \frac{1}{2} q^{\frac{N-1}{2}} \log^c q$. Then

$$a(L) \leq 2a_i q^{\frac{N-1}{2}} \log^{c_1} q \leq \frac{|L|}{\log^{100} q},$$

and by Lemma 6.4 with very high probability

$$||L'| - \mathbb{E}(|L'|)| \leq (a(L)|L|)^{1/2} \log^5 q \leq |L| \log^{-45} q.$$

Since $\mathbb{E}(|L'|) = |L|(1 - P_i^u)$ (see Remark 2.2), with very high probability

$$|L|(1 - P_i^u) - |L| \log^{-45} q \leq |L'| \leq |L|(1 - P_i^u) + |L| \log^{-45} q.$$

For the first inequality

$$\begin{aligned} & |L|(1 - P_i^u - \log^{-45} q) \geq \\ & \geq (1 - \log^{-10} q)^i b_i q^N (1 - P_i^u - \log^{-45} q) \geq (1 - \log^{-10} q)^{i+1} b_{i+1} q^N. \end{aligned}$$

For the second inequality

$$\begin{aligned} & |L|(1 - P_i^u + \log^{-45} q) \leq \\ & \leq (1 + \log^{-10} q)^i b_i q^N (1 - P_i^u + \log^{-45} q) \leq (1 + \log^{-10} q)^{i+1} b_{i+1} q^N. \end{aligned}$$

In fact,

$$1 - P_i^u - \log^{-45} q \geq (1 - \log^{-10} q)(1 - P_i^u) \quad \text{and} \quad 1 - P_i^u + \log^{-45} q \leq (1 + \log^{-10} q)(1 - P_i^u)$$

since

$$1 - P_i^u = \frac{b_{i+1}}{b_i} \geq \frac{1}{2}.$$

□

Proposition 8.6. *Property (15) holds.*

Proof. We have to prove that with very high probability

$$|\Omega_{i+1}| \leq (1 + \log^{-10} q)^{i+1} q^N b'_{i+1}.$$

As already observed $b'_i \leq 2b_i$; moreover by Remark 8.1,

$$|\Omega_i(\ell)| \leq 4 \log^{c_1} q.$$

Let $L = \Omega_i$ and $L' = \Omega_{i+1}$, $A(L, j) = \{x \in \Omega_i | (jx) \cap A_i \neq \emptyset\}$. An upper bound for $A(L, j)$ is $\max_\ell |\Omega_i(\ell)| \times |A_i|$, then

$$a(L) \leq 8a_i q^{\frac{N-1}{2}} \log^{c_1} q,$$

since $|A_i| \leq 2i\theta q^{\frac{N-1}{2}}$. By hypothesis the uncovered points are at least $|L| \geq \frac{1}{2} b_i q^N \geq \frac{1}{2} q^{\frac{N-1}{2}} \log^c q$ (if not, the algorithm can stop and we obtain a cap of the desired size $O\left(q^{\frac{N-1}{2}} \log^c q\right)$). Then

$$a(L) \leq 8a_i q^{\frac{N-1}{2}} \log^{c_1} q \leq \frac{|L|}{\log^{100} q},$$

and by Lemma 6.4 with very high probability

$$||L'| - \mathbb{E}(|L'|)| \leq (a(L)|L|)^{1/2} \log^5 q \leq |L| \log^{-45} q.$$

Since $\mathbb{E}(|L'|) \leq |L|(1 - P_i^\ell) = |L| \frac{b'_{i+1}}{b'_i}$, with very high probability

$$\begin{aligned} |L'| &\leq |L|(1 - P_i^\ell) + |L| \log^{-45} q = |L|(1 - P_i^\ell + \log^{-45} q) \leq \\ &\leq (1 + \log^{-10} q)^i q^N b'_i (1 - P_i^\ell + \log^{-45} q) \leq (1 + \log^{-10} q)^{i+1} q^N b'_{i+1}. \end{aligned}$$

In fact

$$b'_i (1 - P_i^\ell + \log^{-45} q) \leq (1 + \log^{-10} q) b'_{i+1},$$

since

$$1 - P_i^\ell \geq 1 - P_i^u \geq \frac{1}{2}.$$

□

References

References

- [1] J. W. P. Hirschfeld, L. Storme, The packing problem in statistics, coding theory and finite projective spaces: update 2001, *Developments in Mathematics* 3 (2001) 201–246.
- [2] E. M. Gabidulin, A. A. Davydov, L. M. Tombak, Linear codes with covering radius 2 and other new covering codes, *IEEE Trans. Inform. Theory* 37 (1991) 219–224.

- [3] A. A. Davydov, M. Giulietti, S. Marcugini, F. Pambianco, New inductive constructions of complete caps in $\text{PG}(n, q)$, q even, *J. Combin. Des.* 18 (3) (2010) 177–201.
- [4] M. Giulietti, Small complete caps in $\text{PG}(n, q)$, q even, *J. Combin. Des.* 15 (5) (2007) 420–436.
- [5] F. Pambianco, L. Storme, Small complete caps in spaces of even characteristic, *J. Combin. Theory Ser. A* 75 (1) (1996) 70–84.
- [6] A. A. Davydov, P. R. J. Östergård, Recursive constructions of complete caps, *J. Statist. Planning Infer.* 95 (1) (2001) 167–173.
- [7] M. Giulietti, Small complete caps in galois affine spaces, *J. Algebraic Combin.* 25 (2) (2007) 149–168.
- [8] M. Giulietti, F. Pasticci, Quasi-perfect linear codes with minimum distance 4, *IEEE Trans. Inform. Theory* 53 (5) (2007) 1928–1935.
- [9] N. Anbar, D. Bartoli, M. Giulietti, I. Platoni, Small complete caps from singular cubics, *J. Combin. Des.* DOI:10.1002/jcd.21366.
- [10] N. Anbar, D. Bartoli, M. Giulietti, I. Platoni, Small complete caps from singular cubics, II, *J. Algebraic Combin.* (2014) DOI: 10.1007/s10801-014-0532-7.
- [11] A. A. Davydov, G. Faina, S. Marcugini, F. Pambianco, On sizes of complete caps in projective spaces $\text{PG}(n, q)$ and arcs in planes $\text{PG}(2, q)$, *J. Geom.* 94 (1-2) (2009) 31–58.
- [12] G. Pellegrino, On complete caps, not ovaloids, in the space $\text{PG}(3, q)$ with q odd, *Rend. Circ. Mat. Palermo* 47 (1998) 141–168.
- [13] D. Bartoli, G. Faina, M. Giulietti, Small complete caps in three-dimensional galois spaces, *Finite Fields Appl.* 24 (2013) 184–191.
- [14] J. H. Kim, V. H. Vu, Small complete arcs in projective planes, *Combinatorica* 23 (2) (2003) 311–363.
- [15] J. H. Kim, V. H. Vu, Concentration of multivariate polynomials and its applications, *Combinatorica* 20 (3) (2000) 417–434.
- [16] R. A. Brualdi, S. Litsyn, V. S. Pless, *Handbook of Coding Theory*, Vol. 1, Eds. Amsterdam, The Netherlands, 1998, Ch. Covering radius, pp. 755–826.
- [17] G. D. Cohen, S. Honkala, I. S. Litsyn, A. C. Lobstein, *Covering Codes*, Elsevier, Amsterdam, The Netherlands, 1997.
- [18] G. D. Cohen, M. G. Karpovsky, H. F. Mattson Jr., J. R. Schatz, Covering radius - survey and recent results, *IEEE Trans. Inf. Theory* 31 (3) (1985) 328–343.

- [19] G. D. Cohen, A. C. Lobstein, N. J. A. Sloane, Further results on the covering radius of codes, *IEEE Trans. Inf. Theory* 32 (5) (1986) 680–694.
- [20] R. A. Brualdi, V. S. Pless, R. M. Wilson, Short codes with a given covering radius, *IEEE Trans. Inf. Theory* 35 (1) (1989) 99–109.
- [21] D. Bartoli, A. A. Davydov, G. Faina, S. Marcugini, F. Pambianco, New upper bounds on the smallest size of a complete arc in the plane $\text{PG}(2, q)$, *J. Geom.* 104 (2013) 11–43.
- [22] D. Bartoli, A. A. Davydov, G. Faina, S. Marcugini, F. Pambianco, New types of estimates for the smallest size of complete arcs in a finite desarguesian projective plane, *J. Geom.* to appear.
- [23] D. Bartoli, A. A. Davydov, G. Faina, A. Kreshchuk, S. Marcugini, F. Pambianco, Two types of upper bounds on the smallest size of a complete arc in the plane $\text{PG}(2, q)$, in: *Seventh International Workshop on Optimal Codes and Related Topics September 6-12, 2013, Albena, Bulgaria*, www.moi.math.bas.bg/oc2013/a3.pdf, 2013.
- [24] D. Bartoli, A. A. Davydov, G. Faina, S. Marcugini, F. Pambianco, On sizes of complete arcs in $\text{PG}(2, q)$, *Discrete Math.* 312 (3) (2012) 680–698.
- [25] D. Bartoli, A. A. Davydov, G. Faina, S. Marcugini, F. Pambianco, New upper bounds on the smallest size of a complete cap in the space $\text{PG}(3, q)$, in: *Seventh International Workshop on Optimal Codes and Related Topics September 6-12, 2013, Albena, Bulgaria*, <http://www.moi.math.bas.bg/oc2013/a4.pdf>, 2013.